



Structural and temporal variation in the genetic diversity of a European collection of spring two-row barley cultivars and utility for association mapping of quantitative traits

Tondelli, Alessandro; Xu, Xin; Moragues, Marc; Sharma, Rajiv; Schnaithmann, Florian; Ingvarlsen, Christina Rønn; Manninen, Outi; Comadran, Jordi; Russell, Joanne; Waugh, Robbie; Schulman, Alan H.; Pillen, Klaus ; Rasmussen, Søren Kjærsgaard; Kilian, Benjamin; Cattivelli, Luigi; Thomas, William T.B.; Flavell, Andrew J.

Published in:
The Plant Genome

DOI:
[10.3835/plantgenome2013.03.0007](https://doi.org/10.3835/plantgenome2013.03.0007)

Publication date:
2013

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Tondelli, A., Xu, X., Moragues, M., Sharma, R., Schnaithmann, F., Ingvarlsen, C. R., Manninen, O., Comadran, J., Russell, J., Waugh, R., Schulman, A. H., Pillen, K., Rasmussen, S. K., Kilian, B., Cattivelli, L., Thomas, W. T. B., & Flavell, A. J. (2013). Structural and temporal variation in the genetic diversity of a European collection of spring two-row barley cultivars and utility for association mapping of quantitative traits. *The Plant Genome*, 6(2). <https://doi.org/10.3835/plantgenome2013.03.0007>

Structural and Temporal Variation in Genetic Diversity of European Spring Two-Row Barley Cultivars and Association Mapping of Quantitative Traits

Alessandro Tondelli, Xin Xu, Marc Moragues, Rajiv Sharma, Florian Schnaithmann, Christina Ingvaridsen, Outi Manninen, Jordi Comadran, Joanne Russell, Robbie Waugh, Alan H. Schulman, Klaus Pillen, Søren K. Rasmussen, Benjamin Kilian, Luigi Cattivelli, William T. B. Thomas, and Andrew J. Flavell*

Abstract

Two hundred sixteen barley (*Hordeum vulgare* L.) cultivars were selected to represent the diversity and history of European spring two-row barley breeding and to search for alleles controlling agronomic traits by association genetics. The germplasm was genotyped with 7864 gene-based single nucleotide polymorphism markers and corresponding field trial trait data relating to growth and straw strength were obtained at multiple European sites. Analysis of the marker data by statistical population genetics approaches revealed two important trends in the genetic diversity of European two-row spring barley, namely, i) directional selection for approximately 14% of total genetic variation of the population in the last approximately 50 yr and ii) highly uneven genomic distribution of genetic diversity. Association analysis of the phenotypic and genotypic data identified multiple loci affecting the traits investigated, some of which co-map with selected regions. Collectively, these data show that the genetic makeup of European two-row spring barley is evolving under breeder selection, with signs of extinction of diversity in some genomic regions, suggesting that “breeding the best with the best” is leading towards fixation of some breeder targets. Nevertheless, modern germplasm also retains many regions of high diversity, suggesting that site-specific genetic approaches for allele identification and crop improvement such as association genetics are likely to be successful.

CULTIVATED CROP GERMPLASM is a subset of the total genetic diversity originally available in the wild progenitor species. For the Triticeae cereal crops cultivated in Europe (wheat [*Triticum aestivum* L.], barley [*Hordeum vulgare* L.], and rye [*Secale cereale* L.]), the structure of modern cultivar germplasm developed in three phases. First, domestication of the wild plant progenitors occurred around 8,000 to 12,000 yr ago in the Fertile Crescent. Second, locally adapted landraces emerged across Eurasia. Third, modern crop breeding commenced around 1900 (Cleveland and Soleri, 2002). Modern breeding has been reported to have had different effects on the pattern of genetic diversity of Triticeae cereals, varying from substantial reduction in diversity to no significant effect

A. Tondelli and L. Cattivelli, Consiglio per la Ricerca e la sperimentazione in Agricoltura: Genomics Res. Centre, Via S. Protaso, 302, I-29017 Fiorenzuola d'Arda PC, Italy; X. Xu, M. Moragues, and A.J. Flavell, Div. of Plant Sci., Univ. of Dundee at JHI, Invergowrie, Dundee, DD2 5DA, UK; R. Sharma and B. Kilian, Leibniz Inst. of Plant Genet. and Crop Plant Res. (IPK), Corrensstrasse 3, 06466 Gatersleben, Germany; F. Schnaithmann and K. Pillen, Martin-Luther-Universität Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle/Saale, Germany; C. Ingvaridsen and S.K. Rasmussen, Dep. of Plant and Env. Sci., Faculty of Life Sci., Univ. of Copenhagen, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Denmark; O. Manninen and A.H. Schulman, Plant Genomics, Dep. of Biotech. and Food Res., MTT Agrifood Res. Finland, FIN-31600 Jokioinen, Finland; J. Comadran, J. Russell, R. Waugh, and W.T.B. Thomas, The James Hutton Inst., Invergowrie, Dundee, DD2 5DA, UK. Received 20 Mar. 2013. *Corresponding author (a.j.flavell@dundee.ac.uk).

Published in The Plant Genome 6
doi: 10.3835/plantgenome2013.03.0007
© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
An open-access publication

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

Abbreviations: AMOVA, analysis of molecular variance; BLUE, best linear unbiased estimate; CRA, Genomics Research Centre, Fiorenzuola d'Arda, Italy; F_{ST} , Wright's F_{ST} statistic; GWAS, genomewide association scan/scanning; HVCC, *Hordeum vulgare* spring two-row cultivar collection; IPK, Leibniz Institute, Gatersleben, Germany; JHI, James Hutton Institute, Dundee, UK; LD, linkage disequilibrium; MAF, minor allele frequency; PC, principal coordinate; PCO, principal coordinate analysis; PIC, polymorphism information content; QTL, quantitative trait loci; SNP, single nucleotide polymorphism; SSR, simple sequence repeat; UHA, Halle University, Germany.

(Donini et al., 2000; Russell et al., 2000, 2004; Christiansen et al., 2002; Koebner et al., 2003; Maccaferri et al., 2003; Khlestkina et al., 2004; Ordon et al., 2005; Kilian et al., 2006, 2007; Malysheva-Otto et al., 2007). A primary focus of previous studies has been the issue of genetic erosion over time, typically expressed as the number of available alleles per locus, and less attention has been given to the structuring of this diversity (allele combinations and haplotypes) over both time (year of introduction), space (country of origin), and genomic location (Koebner et al., 2003; Ovesná et al., 2013).

Potential problems with the analysis of temporal and geographical trends in germplasm diversity can derive from the sample sets analyzed. If germplasm sets are chosen that lack a significant proportion of the total alleles available in the corresponding total gene pool, then the conclusions may be inaccurate. For example, “new alleles” that have appeared in European agriculture since the 1930s may in fact be present in European landraces and old cultivars that are underrepresented in the analysis (Russell et al., 2000; Malysheva-Otto et al., 2007). To address these issues properly it is important to include a broad representation of European germplasm both new and old.

Another potential problem relates to the choice of markers used to discriminate diversity. If a large number of broadly distributed molecular markers are scored, then markers closely linked to target gene alleles will be available. Many studies of molecular diversity in barley and wheat published to date have been limited by both the nature and quantity of molecular markers. Simple sequence repeat (SSR) markers have been popular in population genetics because they typically capture multiple alleles at each assayed locus, thus providing valuable information for allelic richness in populations. However, they are time consuming both to develop and to score and typically have been deployed in numbers below 50 (e.g., Donini et al., 2000; Russell et al., 2000, 2004; Christiansen et al., 2002; Koebner et al., 2003; Maccaferri et al., 2003; Khlestkina et al., 2004; Ordon et al., 2005; Malysheva-Otto et al., 2007). The latter limitation poses no problems for overall diversity estimation but excludes their use for applications requiring high positional resolution, such as association scanning or searches for signatures of genomic selection.

In recent years the next generation DNA sequencing revolution has facilitated the acquisition of genomewide, high throughput molecular marker platforms for most prominent crop species (e.g., Gupta et al., 2008; Davey et al., 2011). Molecular marker resources for barley—mostly gene-based single nucleotide polymorphism (SNP)—have grown from a few hundred in 2000 to 7864 in 2011 (Rostoks et al., 2006; Close et al., 2009; Comadran et al., 2012). The availability of such marker tools has facilitated the characterization of relatively large germplasm collections (typically hundreds of lines) with genotypic information of high resolution and quality. This has in turn opened the way to both detailed analysis of the distribution of genetic diversity across the genome and association-based searches

for loci tightly linked to traits of interest that segregate in populations (Russell et al., 2011; Kilian and Graner, 2012). Association genetics approaches are commonly used in human and other mammalian systems to identify gene loci encoding traits important to human welfare (Rosenberg et al., 2010). Recently, association studies have successfully identified major genes in barley encoding important agricultural traits (Cockram et al., 2010; Ramsay et al., 2011; Comadran et al., 2012).

Association mapping is prone to high levels of artifact, typically false positives, which are due to nonrandom associations between genotypes in the population studied. Such spurious associations result from strong familial relationships within the total population (i.e., strong population substructure). Optimally, association mapping should be performed on populations with little or no genetic substructure but this approach is feasible even in strongly structured germplasm, provided the nature of that structure is well defined and corrected in the analysis (Comadran et al., 2009, 2011; Pasam et al., 2012; Wang et al., 2012).

Single nucleotide polymorphism-based marker analysis of barley reveals strong substructuring of its genetic variation into four well-defined subpopulations according to the two principal agronomic classes for barley, namely, seasonal growth habit (winter or spring) and ear row number (two- or six-row) (e.g., Cockram et al., 2010). Spring two-row barley represents the largest of these populations and has been the subject of the highest intensity of selective breeding for malting quality and yield. For this study we chose a broad set of cultivated spring two-row barley lines to represent as much as possible the complete diversity and history of European breeding for this crop type in the 20th century. We have investigated the genetic diversity of this collection using a recently developed set of 7864 segregating SNP markers derived from expressed genes (Comadran et al., 2012). Combining these two resources together has allowed us to investigate trends in the structure of genetic diversity in European cultivated barley with respect to time, space, and genomic position. Lastly, we have tested the utility of the collection for genomewide association scanning (GWAS) to discover quantitative loci specifying plant height and straw strength in barley cultivars.

Materials and Methods

Germplasm Assembly and DNA Preparation

Two hundred sixteen cultivated accessions were selected for a *Hordeum vulgare* spring two-row cultivar collection (HVCC₂₁₆) (Supplemental Table S1). Modern European lines (from the United Kingdom, France, Germany, Scandinavia, Czech Republic, Poland, Latvia, Italy, Spain, and Turkey) were sought from the authors' institute stocks and corresponding breeders from the same countries and/or regions. Old lines were identified by pedigree searching in the modern lines and the germplasm was obtained mainly from the stocks of the James Hutton Institute, Dundee, UK (JHI), Leibniz Institute, Gatersleben, Germany

(IPK), and John Innes Cereals Collection. For each line, a single plant was selfed to ensure genetic purity, one of the resulting seed was grown, and DNA was extracted from young leaf tissue, using either Qiagen DNeasy 96 or Tepnel Nucleplex plant DNA extraction kits, according to manufacturers' instructions (Qiagen, Hilden, Germany, or Tepnel Life Sciences PLC, Manchester, UK). The rest of each seed sample was then multiplied to increase stocks for field trial phenotyping experiments.

Genotyping and Linkage Disequilibrium Analysis

A set of 7864 high-confidence, gene-based SNPs, incorporated into a single Illumina iSelect assay (Illumina Inc.), were used to genotype DNAs (Comadran et al., 2012). All genotyping assays were conducted by TraitGenetics GmbH, Gatersleben, Germany. Three thousand eight hundred two of the markers were selected for this study on the basis of both having genetic map positions (Comadran et al., 2012) and segregating in HVCC₂₁₆ with <5% missing score data. A further 1411 segregating iSelect SNPs with <5% missing data that co-map to the genetically mapped iSelect marker set by linkage disequilibrium (LD) (data not shown) were also used, giving 5213 SNPs in total. This marker set was used for the analysis of population structure, polymorphism information content (PIC), and Wright's F_{ST} statistic (F_{ST}) (see below) and a subset of 3989 markers with minor allele frequencies (MAFs) >10% were used for LD and genomewide association analysis.

Pairwise measures of LD (r^2) between genetically mapped SNP markers for each chromosome (Comadran et al., 2012) were calculated with Haploview 4.01 (Barrett et al., 2005). Linkage disequilibrium decay features were evaluated by plotting r^2 values as a function of genetic distances between each SNP pair and fitting a locally weighted scatterplot smoothing (LOESS) curve on the graph.

Population Structure Analysis

Population structure of the HVCC₂₁₆ was investigated by principal coordinate analysis (PCO) based on simple matching of SNP alleles, calculated using Genstat version 14 (VSN International, 2011). Genetic stratification within the collection was also assessed by applying STRUCTURE 2.3 (Pritchard et al., 2000, 2010). STRUCTURE runs were performed for K values from 1 to 10 (three independent runs per K value), using the default setting of the admixture model for the ancestry of individuals and correlated allele frequencies. Burn-in and Markov Chain Monte Carlo values were set to 20,000 and 10,000, respectively (Pritchard et al., 2010), and the best K value was defined by inspection of the PCO plots and by the ΔK method of Evanno et al. (2005). Analysis of molecular variance (AMOVA) and genomewide estimation of population differentiation using Wright's F_{ST} parameter were performed with Arlequin 3.5 (Excoffier et al., 2005). To increase visual resolution, F_{ST} values were transformed to F_{ST}^{10} ($F_{ST} [\exp(10)]$). Polymorphism information content values were calculated for each segregating marker with a known map position (Botstein

et al., 1980; Rostoks et al., 2006) and PIC estimates were plotted with a sliding window of 20 consecutive values stepped by each SNP (Rostoks et al., 2006).

Field Trials and Trait Measurements

The HVCC₂₁₆ was grown in field trials at Genomics Research Centre, Fiorenzuola d'Arda, Italy (CRA), JHI, IPK, and Halle University, Germany (UHA), during 2009 and 2010 (Table 1). Trial data from 213 lines are included in this study (Supplemental Table S1), because three lines were mistakenly omitted. Each of the trials was grown in a two-replicate row and column design with additional filler entries to complete a rectangular grid where necessary. Plots between 2 and 3 m² were grown according to local management practices for sowing rate and chemical inputs. A broad range of traits was scored, with the following analyzed here: plant height (cm) from soil to the bottom of the spike, lodging (percent plot <45° from soil), leaning (percent plot <90° from soil), brackling (percent straw break in the plot), and necking (percent kinked peduncles in the plot).

Trait Data Analysis

All trait analyses used the REML directive in Genstat version 14 (VSN International, 2011). To correct for spatial effects in each of the 34 site and year field trials (Table 1), the most parsimonious model from all possible combinations of random and correlated rows and columns of the trial design was identified by testing for a significant reduction in deviance compared to a randomized complete block model. Best linear unbiased estimates (BLUEs) of mean trait values for each combination of cultivar, trait, site, and year (i.e., 213 BLUEs per trial and trait combination) were then derived for downstream association analysis using the corresponding selected model. In parallel with the above analysis, the relative magnitudes of the variance components genotype, site, year, and combinations thereof were estimated from the raw plot data for each trait over all site and year combinations for which the trait was scored.

Association Analyses

Best linear unbiased estimates (213 lines × 6 or 8 site and year combinations) (Table 1) were analyzed together in quantitative trait loci (QTL) × environment genomewide association scans using the "Single Trait Association Analysis (Multiple Environments)" option of the QTL mapping procedures implemented in Genstat version 14 (VSN International, 2011), with the term environment referring to the site-year combination. The Eigen analysis option was used to control population substructure (Varshney et al., 2012). Using the approach of Boer et al. (2007), we searched for the model that best described the relationships between the genetic variances and covariances over the site-year combinations for each trait and found that the uniform covariance and heterogeneous variance model was the most appropriate. A threshold of $-\log_{10}(P\text{-value}) = 4$ was used to identify significant associations, because it is more stringent than a 5% false

Table 1. Average phenotypic values measured in the spring two-row barley lines and in the different STRUCTURE (Pritchard et al., 2000, 2010) groups.

| Trait | STRUCTURE group | U.K. (Dundee) | | D [†] (Gatersleben) | | D (Halle) | | I [‡] (Fiorenzuola) | |
|-------------------|-----------------|--------------------|--------------------|------------------------------|--------------------|--------------------|--------------------|------------------------------|--------------------|
| | | JHI09 [§] | JHI10 [§] | IPK09 [§] | IPK10 [§] | UHA09 [§] | UHA10 [§] | CRA09 [§] | CRA10 [§] |
| Plant height, cm | Overall | 90.9*** | 87.1*** | 90.8*** | 75.3*** | 76.4*** | 83.3*** | 65.6*** | 52.8*** |
| | K1 (modern) | 80.8 | 79.1 | 82.4 | 69.1 | 72.8 | 79.4 | 63.1 | 50.5 |
| | Admixed | 92.3 | 89.3 | 92.2 | 76.3 | 78.2 | 83.9 | 66.1 | 53.4 |
| | K2 (old) | 100.0 | 92.7 | 98.3 | 79.9 | 78.3 | 86.3 | 67.3 | 54.0 |
| Lodging, % plot | Overall | 0.8* | 3.7** | 3.4*** | 6.1*** | 16.2*** | 0.1 | 2.6** | 0.7* |
| | K1 (modern) | 0.0 | 0.8 | 0.0 | 2.5 | 2.4 | 0.2 | 0.6 | 0.1 |
| | Admixed | 0.2 | 2.6 | 1.4 | 2.9 | 18.9 | 0.0 | 3.0 | 1.0 |
| | K2 (old) | 3.1 | 8.5 | 9.8 | 12.0 | 29.3 | 0.0 | 3.7 | 1.0 |
| Leaning, % plot | Overall | 11.6*** | 8.4*** | 7.0*** | 9.2*** | 34.0*** | 5.5*** | na [¶] | na |
| | K1 (modern) | 1.1 | 2.3 | 1.2 | 4.2 | 19.8 | 0.5 | na | na |
| | Admixed | 11.6 | 6.6 | 4.5 | 4.6 | 38.3 | 5.7 | na | na |
| | K2 (old) | 20.1 | 15.9 | 16.5 | 17.9 | 45.4 | 9.8 | na | na |
| Brackling, % plot | Overall | 19.7 | 4.5** | 2.1** | 5.9* | 33.8*** | 5.3*** | na | na |
| | K1 (modern) | 22.6 | 1.5 | 0.0 | 3.3 | 21.8 | 1.2 | na | na |
| | Admixed | 23.9 | 4.0 | 1.3 | 7.5 | 38.2 | 4.9 | na | na |
| | K2 (old) | 23.9 | 7.3 | 5.5 | 8.0 | 42.2 | 9.3 | na | na |
| Necking, % plot | Overall | 45.2** | 8.8 | 1.4 | 8.9** | 0.2 | 0.9 | na | na |
| | K1 (modern) | 34.3 | 8.0 | 1.5 | 3.7 | 0.1 | 0.4 | na | na |
| | Admixed | 53.3 | 10.0 | 1.0 | 11.9 | 0.0 | 1.1 | na | na |
| | K2 (old) | 50.1 | 8.6 | 2.1 | 11.4 | 0.6 | 1.7 | na | na |

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

[†]D, Germany.

[‡]I = Italy.

[§]JHI, James Hutton Institute, Dundee, UK; IPK, Leibniz Institute, Gatersleben, Germany; UHA, Halle University, Germany; CRA, Genomics Research Centre, Fiorenzuola d'Arda, Italy. Columns ending "09" and "10" refer to trials carried out in 2009 and 2010, respectively.

[¶]na, not analyzed.

discovery rate for the traits height, lodging, and necking. Estimated allelic effects were expressed relative to the performance of cultivar Optic, to show their directionalities (Table 2). Confidence intervals for the GWAS peaks were defined by the SNP with the highest $-\log_{10}[P\text{-value}]$ within a ± 10 cM window and extended to the most distant marker within that window whose $-\log_{10}[P\text{-value}]$ was within a value of 1 of the peak value.

Results

Assembly and Genotyping of a European Spring Barley Two-Row Cultivar Collection

Two hundred sixteen *H. vulgare* cultivar accessions (HVCC₂₁₆) were selected to represent the diversity and evolution of European cultivated two-row spring barley in the 20th century (Materials and Methods; Supplemental Table S1). The lines were genotyped using an Illumina iSelect marker set (Comadran et al., 2012) comprising 7864 independent SNPs, 3989 of which were suitable for LD evaluation, on the basis of marker quality, allele frequency, and genetic mapability (Materials and Methods). Intrachromosomal LD decay (r^2) among these markers for the HVCC₂₁₆ varies between 0.048 and 0.099 for the seven

barley chromosomes (Supplemental Fig. S1A) and average LD drops below a threshold r^2 value of 0.15 (Comadran et al., 2009) after 4 to 6 cM (Supplemental Fig. S1B).

A Temporal Trend in the Diversity of Spring Two-Row Barley

Trends in the genetic stratification of the HVCC₂₁₆ were investigated using both the Bayesian approach STRUCTURE (Pritchard et al., 2000; Evanno et al., 2005), which has been used successfully for barley and similar inbreeding crop situations (e.g., Jing et al., 2010; Comadran et al., 2011), and PCO (Fig. 1; Supplemental Fig. S2). For this analysis the 3989 markers used above were supplemented by a further 1224 segregating SNPs with MAF <10%, giving 5213 in total (Materials and Methods). STRUCTURE analysis indicated the presence of two groups, K1 and K2, of 73 and 71 accessions, respectively, with the remaining 72 barley cultivars being classified as admixed (Fig. 1A). Superimposing the STRUCTURE grouping on the results from the PCO showed that K1 or K2 assignment corresponds to subdivision of the population along the first principal coordinate (PC1) (Fig. 1B). To test this correspondence statistically, we performed AMOVA analysis. A significant portion of the genotypic differentiation (14.1%)

Table 2. Results of the genomewide association scans carried out on the spring two-row barleys.

| Character | Marker [†] | Chromosome | Genetic map position (cM) [‡] | log ₁₀ (P-value) [§] | Corresponding cultivar Optic data | | | |
|------------------|------------------------|------------|--|--|-----------------------------------|----------------------|------------------------|------------------------|
| | | | | | Optic allele frequency | Average Optic effect | Increase | Decrease |
| Height | i_SCRI_RS_148560 | 1H | 13.39 | 5.613 | 0.746 | 1.65 | 6 | 0 |
| Lodging | i_11_10259 | 1H | 46.32 | 4.959 | 0.845 | −2.3 | 0 | 3 |
| <i>Brackling</i> | <i>i_11_20810</i> | <i>1H</i> | <i>48.58</i> | <i>3.928</i> | <i>0.883</i> | <i>−2.31</i> | <i>1</i> | <i>2</i> |
| Height | i_12_30110 | 1H | 48.94 | 5.65 | 0.845 | 3.93 | 6 | 0 |
| Height | i_SCRI_RS_170110 | 1H | 116.78 | 5.928 | 0.897 | −3.79 | 0 | 6 |
| Height | i_SCRI_RS_206337 | 1H | 130.81 | 5.628 | 0.155 | −2.88 | 0 | 6 |
| Height | i_SCRI_RS_219333 | 2H | 0 | 5.582 | 0.803 | −1.74 | 0 | 5 |
| Leaning | i_SCRI_RS_219333 | 2H | 0 | 4.131 | 0.803 | −4.99 | 0 | 5 |
| Lodging | i_SCRI_RS_219333 | 2H | 0 | 7.207 | 0.803 | −2.47 | 0 | 4 |
| Height | i_SCRI_RS_174935 | 2H | 37.82 | 8.000 | 0.859 | −2.83 | 0 | 6 |
| Height | i_12_21288 | 2H | 56.37 | 7.89 | 0.85 | −2.53 | 0 | 6 |
| Lodging | i_SCRI_RS_188579 | 2H | 135.62 | 4.172 | 0.896 | −2.95 | 0 | 4 |
| Necking | i_11_20797 | 3H | 1.63 | 5.468 | 0.751 | −2.76 | 0 | 2 |
| Height | i_SCRI_RS_130264 | 3H | 8.78 | 4.897 | 0.592 | 1.51 | 5 | 0 |
| Lodging | i_11_10926 | 3H | 51.2 | 7.443 | 0.789 | 3.72 | 4 | 0 |
| Height | i_11_10456 | 3H | 51.35 | 19.887 | 0.249 | −8.38 | 0 | 7 |
| Height | i_SCRI_RS_146429 | 3H | 83.07 | 4.242 | 0.516 | −1.41 | 0 | 4 |
| Necking | i_SCRI_RS_206510 | 3H | 103.75 | 4.100 | 0.878 | −3.07 | 0 | 2 |
| Height | i_11_10754 | 3H | 108.57 | 6.628 | 0.82 | −2.64 | 0 | 6 |
| Lodging | i_SCRI_RS_120973 | 3H | 108.57 | 5.265 | 0.812 | −1.98 | 0 | 3 |
| Height | i_SCRI_RS_180847 | 3H | 145.33 | 4.072 | 0.554 | −1.56 | 0 | 6 |
| Height | i_12_30839 | 4H | 54.04 | 4.484 | 0.869 | −2.82 | 0 | 7 |
| <i>Necking</i> | <i>i_SCRI_RS_22172</i> | <i>4H</i> | <i>54.04</i> | <i>3.873</i> | <i>0.836</i> | <i>0.35</i> | <i>0</i> | <i>0</i> |
| Lodging | i_SCRI_RS_203585 | 5H | 30.07 | 4.467 | 0.366 | 1.74 | 3 | 0 |
| Lodging | i_SCRI_RS_49958 | 5H | 52.3 | 4.293 | 0.776 | −1.53 | 0 | 1 |
| Necking | i_11_11473 | 5H | 81.32 | 7.779 | 0.897 | −4.36 | 0 | 3 |
| Lodging | i_SCRI_RS_162696 | 5H | 113.89 | 5.556 | 0.897 | −2.22 | 0 | 4 |
| Height | i_11_21422 | 5H | 120.35 | 8.089 | 0.606 | −2.4 | 0 | 6 |
| Necking | i_SCRI_RS_138933 | 5H | 125.63 | 5.457 | 0.596 | 3.55 | 2 | 0 |
| Necking | i_SCRI_RS_720 | 5H | 146.11 | 4.563 | 0.493 | 2.68 | 2 | 0 |
| Lodging | i_SCRI_RS_138529 | 6H | 55.38 | 4.471 | 0.718 | −3.09 | 0 | 3 |
| Height | i_11_20468 | 6H | 66.78 | 4.803 | 0.869 | −2.33 | 0 | 5 |
| Height | i_SCRI_RS_166536 | 7H | 40.37 | 5.014 | 0.775 | 2.51 | 7 | 0 |
| Height | i_SCRI_RS_154003 | 7H | 121.81 | 12.656 | 0.788 | 3.36 | 7 | 0 |

[†]Comadran et al. (2012).[‡]Marker with highest $-\log_{10}$ [P-value] score within confidence interval (see Fig. 3).[§]Significance scores (see Materials and Methods). Two genomewide association scanning peaks discussed in the text with $-\log_{10}$ [P-value] scores just below the threshold of 4 are in italics.^{||}Number of environments (out of eight for plant height or six for the other traits) where the Optic allele showed a positive or negative effect on the trait.

was found between the K1, K2, and admixed groupings, with the remaining 85.9% being retained within them. After removing the admixed accessions, the between-group component of molecular variance increased to 24.2%. Looking for possible causes of this stratification, we observed a correlation between the year of release for the different cultivars and the K1 or K2 group classification ($r^2 = 0.608$, $P < 0.0001$), with K1 and K2 groups corresponding respectively to mainly modern and old barley lines (Fig. 1C; Supplemental Table S1). No correlation with country of origin was apparent (data not shown).

Genomic regions involved in differentiating K1 and K2 genotypes were revealed by plotting marker-associated F_{ST} values for all mapped SNPs (Fig. 2A).

Signatures of diversification (high F_{ST} regions nearing fixation of opposing alleles in the two populations) were observed at several locations on chromosome 1H, close to the telomeres of chromosomes 2HS (8.2 cM) and 3HL (155.0 cM), and at two positions on chromosome 5H (44.2 and 131.2 cM) (Fig. 2A). To highlight differences in the patterns of polymorphism between the K1 and K2 STRUCTURE (Pritchard et al., 2000, 2010) groups along the barley chromosomes, PIC values were calculated for the same mapped SNP markers (Fig. 2B). The K1 and K2 groups show overlapping PIC profiles, with some low polymorphism regions shared (e.g., the centromeric region of chromosome 2H) and others specific to one group. This indicates that some genomic regions have

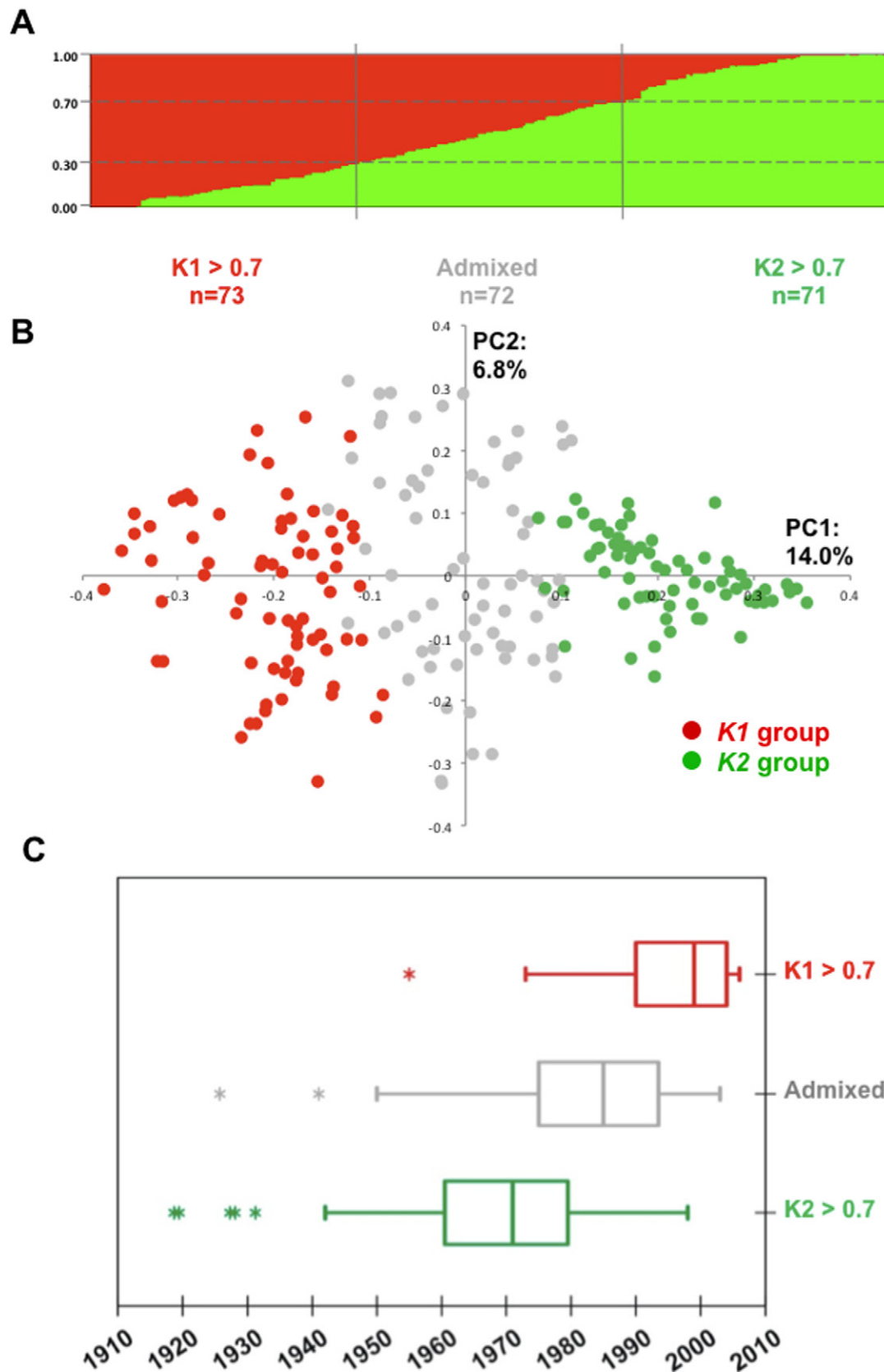


Figure 1. Genetic stratification of spring two-row barley accessions. (A) STRUcTURE (Pritchard et al., 2000, 2010) output for $K = 2$. (B) Principal coordinate analysis plot colored according to the STRUcTURE grouping in (A). (C) Box plot comparing the year of cultivar release, according to the STRUcTURE grouping in (A). Boxes represent the interquartile ranges (IQRs) and include data falling between the 25th and the 75th percentiles; flanking lines and asterisks mark data within $1.5 \times \text{IQR}$ and $3 \times \text{IQR}$, respectively. PC, principal coordinate.

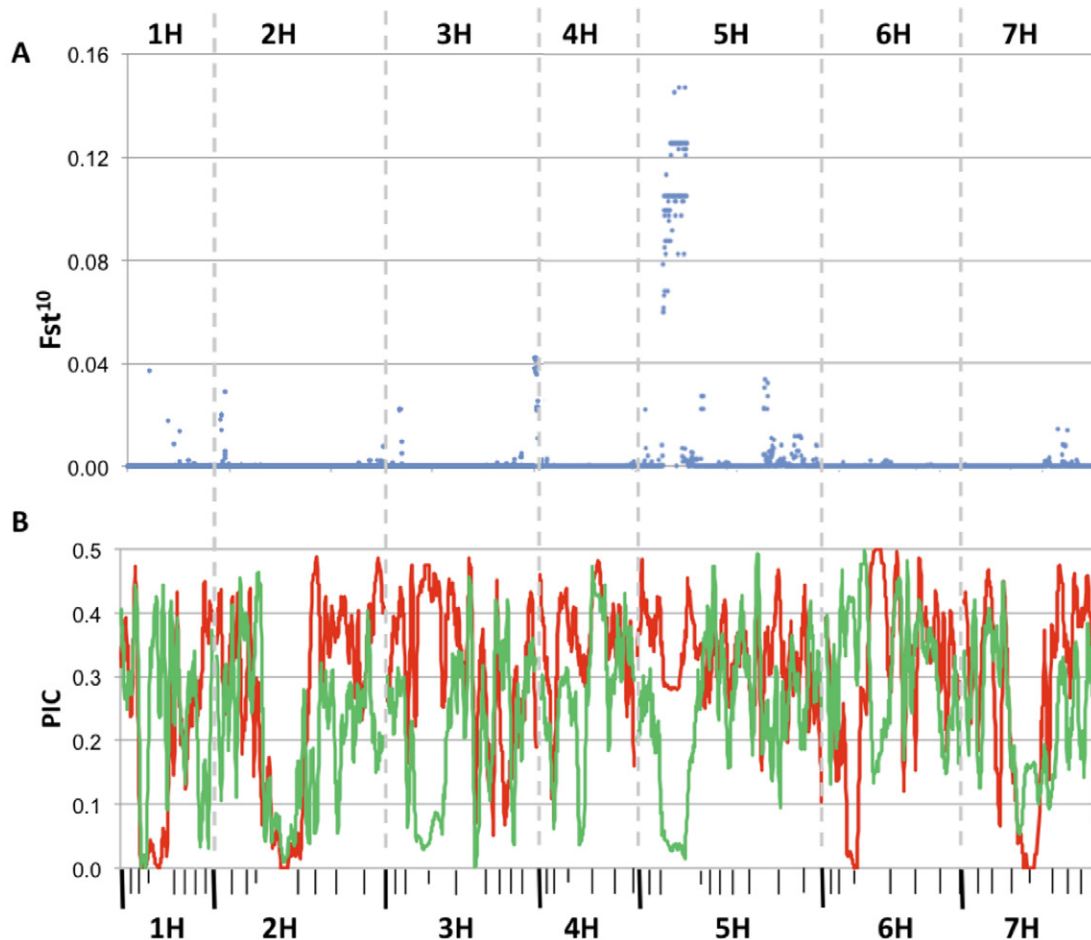


Figure 2. Genomewide profiles of genetic differentiation between STRUCTURE (Pritchard et al., 2000, 2010) groups of spring two-row barley. (A) Transformed F_{ST}^{10} ($F_{ST} [\exp(10)]$) value (see Materials and Methods) of each marker derived from K1 or K2 STRUCTURE group occupancy (Fig. 1) and (B) genomewide profiles of polymorphism information content (PIC) of markers for the same STRUCTURE groups are plotted against genomic position. The seven barley chromosomes 1H through 7H are delineated by vertical dotted lines and marker positions correspond between (A) and (B). Polymorphism information content values are given as rolling averages based on 20 adjacent loci. Single nucleotide polymorphisms are scaled by map order (Comadran et al., 2012), and corresponding genetic map positions are indicated at 20 cM intervals by vertical bars at the bottom of the figure, with each median 10 cM position for the seven pericentromeric 20 cM intervals indicated by a shorter bar.

remained fixed during the past approximately 60 yr of barley breeding whereas others have either approached fixation or escaped from it. A prominent example of the former is the centromeric region of chromosome 2H, which contains the *HvCEN* major gene locus at which an allele conferring late flowering, with an associated large yield increase, is fixed in spring barley cultivars (Comadran et al., 2012) (Fig. 2B). Genomic regions with prominent low PIC values for modern two-row spring barley alone are apparent on chromosomes 1H (47.8–55.4 cM), 6H (30.2–53.6 cM), and 7H (29.8–47.6 cM) (Fig. 2B). Such regional drops in genetic diversity presumably derive from human selection and reflect the targets of modern barley breeding. For example, several QTL and candidate genes (e.g., *HvCslF9* and *HvGlb1*) for malting quality traits have been mapped to the low PIC value pericentromeric region of chromosome 1H (47.5–50.9 cM) (Burton et al., 2008; Laidò et al., 2009). We also identified several regions with lower PIC values in old

cultivars on all chromosomes. On average, slightly lower average PIC values were found for K2 (old) cultivars (0.24) than for K1 (modern) European barley lines (0.30). Possible reasons for this are discussed later.

Genomewide Association Scanning in the Spring Two-Row Barley Population

To explore the utility of the population for GWAS, the HVCC₂₁₆ was trialed in eight site and year combinations across Europe (see Materials and Methods). Large phenotypic differences were detected in the accessions for all the traits under evaluation (Supplemental Table S2). Variance components analysis showed that the genetic main effect ranged from 4% (necking) to 60% (height) of the total phenotypic variation (Supplemental Fig. S3). Genetic interactions (effects of combinations of genotype with site and/or year) were relatively minor for height and leaning but noticeably greater than the main genetic effect for the other traits studied.

For plant height, the mean values of the K1 (modern) and K2 (old) groups were significantly different ($P < 0.001$) in all the field trial comparisons (Table 1). Modern spring two-row barley lines (K1) were between 3.9 cm (CRA 2010 trial) and 20.4 cm (JHI 2009 trial) shorter on average than old accessions (K2) and admixed lines showed an intermediate phenotype. Unsurprisingly, highly significant differences ($P < 0.001$) were also detected between groups K1 and K2 for plant leaning (see Materials and Methods for definitions) in the six trials analyzed and for plant lodging in the IPK 2009, IPK 2010, and UHA 2009 trials (Table 1). Higher incidences of straw brackling were noticed in old barley genotypes whereas smaller differences were observed for straw necking.

Genomewide association scans were performed across all available site and year environments (see Materials and Methods) (Fig. 3). The strong genotype \times environment interactions described above (Supplemental Fig. S3) indicated that QTL \times environment association analysis was appropriate for this study (Materials and Methods). Most significant associations ($-\log_{10}[P\text{-value}] > 4$) are for plant height (17 regions), with lodging accounting for nine regions, necking accounting for five regions, leaning accounting for one significant GWAS peak, and brackling accounting for none (Table 2). Most of the highly significant GWAS peaks ($-\log_{10}[P\text{-value}] > 6$) are also for plant height (6/9), on chromosomes 2H (37.8 and 56.4 cM), 3H (51.3 and 108.6 cM), 5H (120.3 cM), and 7H (121.8 cM). Two of the regions described above with low PIC value for modern (K1) barley contain significant genotype–phenotype associations for height (on chromosome 1H at 48.9 cM and on chromosome 7H at 40.4 cM) (Fig. 2B and 3) and the first of these is also a significant GWAS peak for lodging and a near-significant peak ($-\log_{10}[P\text{-value}] = 3.928$) for brackling (Table 2). Many of the GWAS peaks co-localize with corresponding peaks for other traits (14/32) (Fig. 3B) and in four cases, three or more peaks co-locate. This is not surprising as these traits are cross-related, by the common factors of straw strength and potential for wind damage via height exposure.

To ascertain their directionalities, genotypic allelic effects were normalized relative to the performance of Optic, which represents a widely used reference spring two-row barley. The Optic peak SNP alleles for height are associated with significantly reduced height in 36% of marker and trial combinations and increases in 33%. The Optic allele at SNP i_11_10754 on chromosome 3H, which is closely linked to *sdw1* (Malosetti et al., 2011), is associated with a height reduction of 2.6 cm and was significant in all environments except the CRA 2009 and CRA 2010 trials. The biggest GWAS region associated with height is also located on chromosome 3H where the Optic allele at SNP i_11_10456 is associated with a reduction of 8.4 cm (Table 2) and was significant in all but one environment (the CRA 2010 trial). This region contains multiple genes with known effect on height, including *uzu* (Pasam et al., 2012) and *ari-a*, *ert-c*, and *ert-ii* (Lundqvist et al., 1996). Optic alleles at five other

QTL (SNPs i_SCRI_RS_170110, i_SCRI_RS_206337, i_SCRI_RS_174935, i_12_21288, and i_12_30839) (Table 2) are associated with significant height reductions > 2.5 cm. The first two of these, on chromosome 1H, may correspond to a QTL in the same region reported by von Korff et al. (2006) and Inostroza et al. (2009) and the rest share approximate genetic map positions with associations described by others (i_SCRI_RS_174935 [Pasam et al., 2012], SNP i_12_21288 [Marquez-Cedillo et al., 2001], and i_12_30839 [Hayes et al., 1993]).

Optic is also associated with significant height increases and for three of these associations they are in excess of 2.5 cm (Table 2) (SNPs i_12_30110, i_SCRI_RS_166536, and i_SCRI_RS_154003). The last is the most significant and the corresponding Optic allele effect is the second largest, at 3.4 cm. This region (on chromosome 7H at 121.8 cM) is coincident with the location of a semidwarfing gene (Yu et al., 2010). Single nucleotide polymorphism i_12_30110 is associated with the largest height increase (3.9 cm) and it maps close to the QHt.HaMo-1H QTL (Marquez-Cedillo et al., 2001). Lastly, SNP i_SCRI_RS_166536 lies close to ABG603 (Zhu et al., 1999).

Significant associations with the straw strength characters lodging, leaning, and particularly brackling and necking were not detected as frequently as those for height (see above). In general, the peak SNP alleles derived from Optic mostly reduce height and lodging but have mixed effects on necking. Most of the height associations noted above are located in regions also affecting straw strength characters. A notable example is the peak at 43 to 49 cM on chromosome 1H, which is also associated with lodging ($-\log_{10}[P\text{-value}] = 4.959$) and brackling ($-\log_{10}(P\text{-value}) = 3.928$) (Fig. 3). While the Optic alleles increase height, they decrease lodging and brackling although none of the peak SNPs coincide and these differences could therefore reflect different recombination events. Similar inconsistencies are seen for height and lodging on chromosome 3H at 51 cM and height and necking on chromosome 4H at 54 cM. In contrast, regions that co-map exactly mostly show predictably linked effects, for example, the telomeric peak on chromosome 2HS (marker i_SCRI_RS_219333 at 0 cM) is associated with reduced height, leaning, and lodging for the Optic allele.

Discussion

Our characterization of the genetic and phenotypic diversity of European cultivated two-row spring barley using a new set of 7864 SNP markers has allowed us to reconsider previous descriptions of the diversity and history of cultivated barley in the 20th century, to explore the distribution of this diversity across the barley genome, and to search for loci underlying quantitative traits important for agriculture.

A Temporal Trend in the Genetic Diversity of Barley

We find a clear temporal trend in the diversity of European two-row spring barleys, corresponding to 14% of its

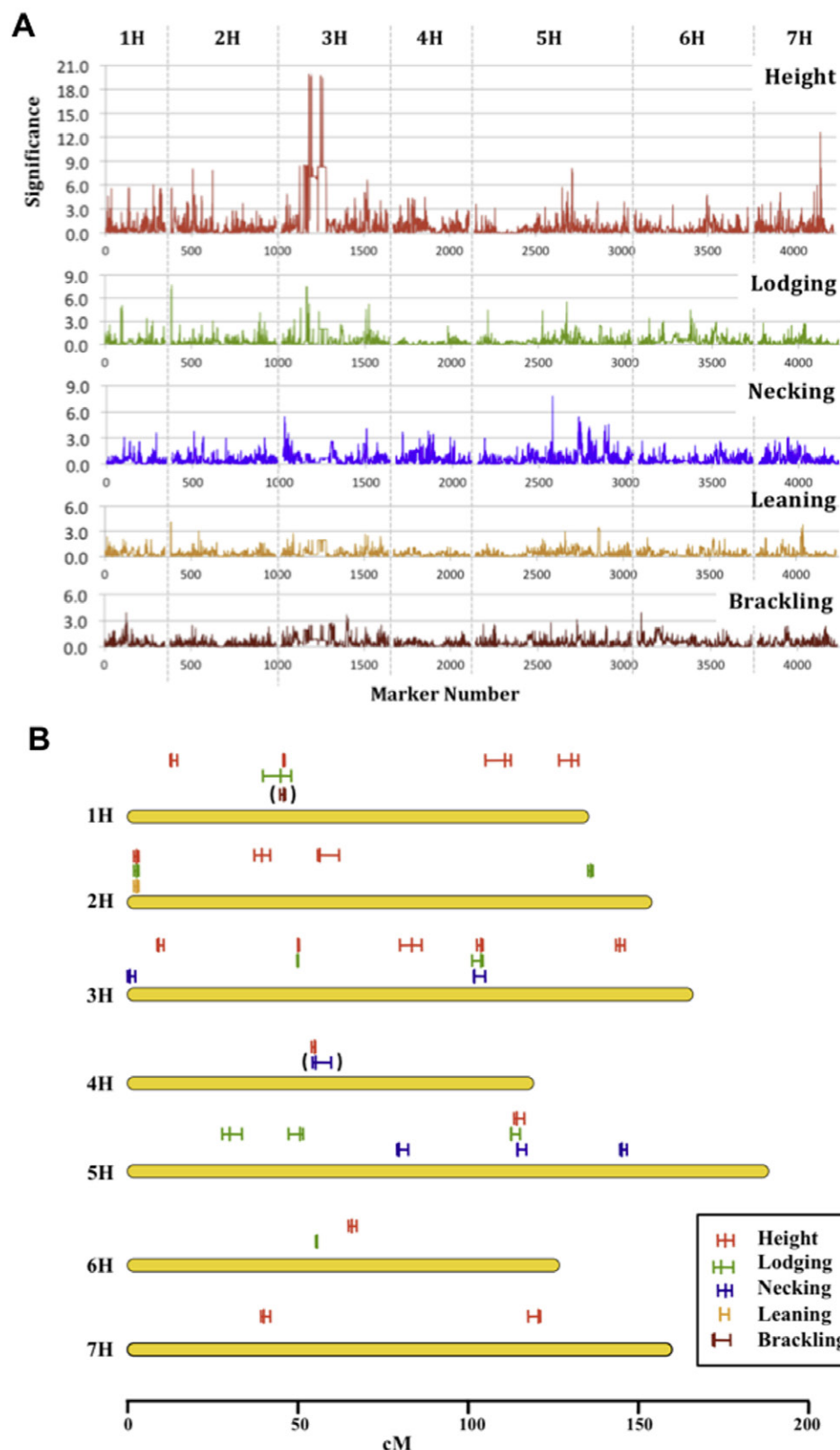


Figure 3. Loci affecting height and straw-strength traits identified by genomewide association scanning (GWAS) in the *Hordeum vulgare* spring two-row cultivar collection. Traits are color coded (see legend). (A) Genomewide association scanning peak traces ($[-\log_{10}(P\text{-value})]$ significance scores) are plotted against map order (Comadran et al., 2012), not cM genetic map position. Chromosomes are separated by vertical dotted lines. (B) Significant GWAS peaks $[-\log_{10}(P\text{-value}) > 4]$, with corresponding confidence intervals indicated by flanking bars, are aligned by genetic map position on the seven barley chromosomes. Data for all significant GWAS peaks are listed in Table 2.

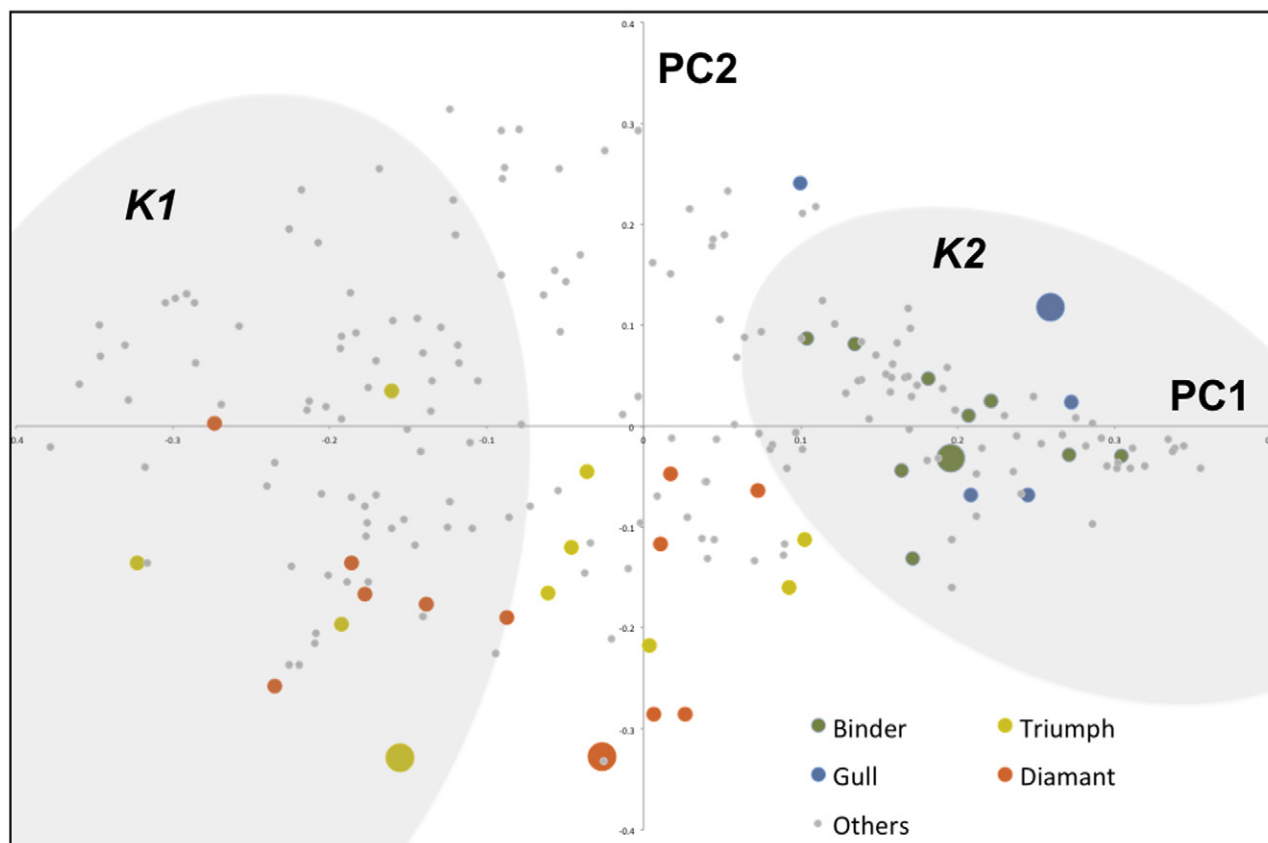


Figure 4. Contributions of particular cultivars to the diversity of European spring two-row barley. The principal coordinate analysis plot from Fig. 1B is shown, with the contributions of four important founder cultivars in the pedigrees of the lines indicated by colored circles. The actual cultivars Binder, Gull, Diamant, and Triumph are shown by larger colored circles. The approximate extents of the K1 and K2 STRUCTURE (Pritchard et al., 2000, 2010) groups are shown by shading (see Fig. 1). PC, principal coordinate.

total genotypic variability (Fig. 1B). Koebner et al. (2003) used PCOs of amplified fragment length polymorphism, SSR, and morphological data, which distinguished pre-1970 from 1990s spring barleys, with 6% of variation observed for the corresponding genetic component. Conversely, Ovesná et al. (2013) observed a shift along the PC1, corresponding to 23% of total variance, in a collection of 94 Czech malting barley cultivars analyzed with 234 DNA Diversity Array Technology (DArT) markers. The different germplasm sets and marker types used in these studies and ours are likely to introduce different biases to the deduced diversity but we suggest that they all reveal a common trend in allele frequency in European barley genetic diversity. In view of its broad distribution, gene-based nature, and high positional resolution, we recommend the barley 7864 iSelect SNP platform as the most appropriate tool for revealing variation in barley genetic diversity.

Our results show that older and more modern barley varieties are distinguishable by the two independent statistical approaches STRUCTURE (Pritchard et al., 2000, 2010) and PCO. We observe a gradient of differentiation along PCO1 (Fig. 1B) that we suggest has resulted from directional selection on specific breeding targets. This reflects the change in European barley breeding after 1970, when cultivar Triumph (also

known as Trumpf [Supplemental Table S1]) emerged as a widespread donor of yield, malting quality, and disease resistance (Fischbeck, 1992). In this view, the K2 (old) group represents old cultivars such as Binder (derived from the Scandinavian ‘Hanna’) and Gull (Sweden), plus their derivatives (Fig. 4), while the K1 (modern) group derives mostly from the same Hanna, ‘Bavaria’, and ‘Danubia’ (both Bavarian), plus intercrosses that led to Triumph via ‘Diamant’. Support for this model comes from the following observations: i) There is a progression along the first coordinate of the PCO plot, leading from Binder and Gull in the K2 group to Diamant in the admixed group and Triumph in the K1 (modern) STRUCTURE group (Fig. 4). Cultivars with Binder and Gull in their pedigrees are in the K2 (old) group, and lines with Diamant and Triumph as immediate ancestors are either in the admixed or K1 groups. ii) *sdw1* types that are derived from ‘Abed Denso’, an independent Scandinavian source of the dwarfing gene (Haahr and von Wettstein, 1976) and cultivars containing introgressed disease resistances fall into the admixed set, and the *sdw1* types derived from Triumph reside in the K1 (modern) group (Supplemental Table S1).

Our data reveal no geographical trends in the diversity of European cultivar barley, in agreement with previous observations (Kraakman et al., 2004;

Malysheva-Otto et al., 2007). This is unsurprising, as European breeders use each other's germplasm enthusiastically in their breeding programs. This artificial outcrossing, coupled with the long recombination history of the germplasm has created a highly diverse germplasm stock without major population subdivisions (Rostoks et al., 2006).

Trends in Intragenomic Diversity

Our analysis of SNP variation at thousands of genetically mapped loci has enabled the exploration of the genomic context of trends in barley diversity (Fig. 2). We find that the major regions differentiating K1 and K2 genotypes of spring two-row barley map to multiple dispersed sites in the barley genome (Fig. 2A), with some of these depleted for genetic diversity (Fig. 2B). Directional selection might be expected to reduce diversity universally but this is not necessarily the case for breeding efforts, because different breeders use different donor germplasms in their programs.

Our data indicate that modern (K1) spring two-row barley has fewer genomic regions with low PIC values than the older K2 barleys. Therefore, modern germplasm apparently has a slightly higher overall genetic diversity than old (Fig. 2B). Earlier efforts variously found either no change (Koebner et al., 2003), loss, or gain of SSR alleles depending on the locus (Malysheva-Otto et al., 2007), or an overall reduction in diversity over time (Russell et al., 2000). Our data suggest that the second of these views is the best representation, with the overall balance moving towards higher total diversity. Again, this might seem unexpected, but barley breeding was already using limited, cross-related parents 100 yr ago, which would lead to regional loss in diversity, and modern breeding began to introduce disease resistance via outcrossing to exotic germplasm from the 1960s. A possible example of this is the centromere of chromosome 5H, which shows high F_{ST} , high K1 PIC values, and low K2 PIC values (Fig. 2), suggesting that different multigene haplotypes are segregating in this region. These haplotypes are strongly skewed in the old (K2) group and segregating in the modern (K1) group. Several leaf rust resistance loci are located in this region (on chromosome 5H at 44.2 cM), including *Rph2* (Borovkova et al., 1997), and breeder-mediated selection pressure for resistance may underlie this selective sweep.

One caveat to our conclusions is that SNP markers can be subject to ascertainment bias, derived from the choice of material for their discovery (Moragues et al., 2010). The SNPs used here were developed from a small number of cultivars that were released mostly after 1980 (Comadran et al., 2012; Supplemental Table S1) and they may underestimate diversity in older cultivars. The chromosome 7H central region has low diversity for a subset of our iSelect marker set, using a 53 sample set that has 45% sample overlap with ours (Rostoks et al., 2006), suggesting that low diversity is universal rather than biased. This interpretation is supported by sequence data for 105 genes on chromosome 7H, among 10 barley

cultivars, four of which are shared with our cultivar set (our unpublished data). The sequence data also shows a pronounced diversity drop (π statistic) in the same region. Thus in this case our markers are revealing a largely unbiased view of the regional barley diversity. A clear resolution of this issue awaits comparative sequencing of the gene components of all the lines studied here.

Association Genetics of Straw-Related Traits in Spring Two-Row Barley Cultivars

Our choice of European spring two-row barley and traits related to growth and straw strength for GWAS was conservative in one way, as the germplasm lacks strong population substructure (Fig. 2B) and thus well suited for GWAS, and optimistic in the other, since genotype effects for some of the traits are modest (Supplemental Fig. S3). The results have been correspondingly successful in some ways and less rewarding in others. We identified 32 significant putative QTL regions for plant height, lodging, and necking (Table 2; Fig. 3), many of which coincide encouragingly with known loci affecting these traits (see Results). In contrast, we identified only one promising peak for each of leaning and brackling, both of which are very close to our chosen significance threshold ($-\log_{10}[P\text{-value}]$) of 4. These results are somewhat surprising because the trait leaning showed a high genotype component in our trials and necking showed the lowest component (Supplemental Fig. S3). Therefore, our success in discovering QTL is not simply related to the corresponding proportions of genetic variation for the traits. We suggest that our correction for population substructure has removed most of the variation for leaning and brackling but this has been less pronounced for the other three traits. In support of this suggestion, there is a marked contrast between the K1 and K2 groups for leaning (Table 1) and this grouping underlies the most prominent component of our genetic variation (Fig. 1B), which was used for population substructure correction.

“Old” (K2) and “modern” (K1) barleys differed significantly for their height (Table 1) in all testing environments, suggesting directional selection for this parameter over the past approximately 50 yr. Furthermore, two of the three regions with low PIC value in the modern K1 group contain GWAS peaks for height (Fig. 2B; Table 2). However, other GWAS peaks for height, including the dominant one on chromosome 3H, do not correspond with obvious regions of directional selection. This may be due to lack of fixation in the modern gene pool; indeed, the low height *Optic* allele at the top SNP for height (SNP_i_11_10456) is carried by just 2% of K2 (old) cultivars and 35% of K1 (modern) group cultivars. We expect the allele frequency for this SNP to increase in the future as directional selection by breeders continues.

The introduction of new diversity, either by changes in allele frequency or introduction of new alleles, would lead to an increase in local PIC value and a GWAS peak. Characteristics such as disease resistance, height, and maturity have been strongly selected by breeders, leading

to corresponding changes in allele frequencies at the underlying loci and their flanking genomic regions in strong LD. Additionally, the complex quantitative nature of the traits studied here complicates naïve interpretation of the connections between selection, diversity, and association. Nevertheless, the co-location of multiple QTL for different traits identified here supports the notion that these particular loci specify common aspects of barley straw physiology (e.g., lignin content, elasticity, root traits) and/or geometry (e.g., height, stem thickness).

Our data also highlight the importance of performing phenotypic trials over different environments. All of the effects reported here are QTL \times environment interactions although most are scaling effects. We have identified several genomic regions with notable height effects that have been and continue to be exploited in European spring barley breeding. The most obvious trend is the use of the Diamant-derived *sdw1* allele, which is widespread in the K1 (modern) group. The height effect that we have detected from GWAS is relatively small compared to that reported from biparental studies (e.g., Thomas et al., 1995) but this trait segregates within PCO1 (see above) and the genetic structure correction within our GWAS has likely absorbed much of the genetic variation for height within this population, as proposed for leaning and brackling above.

The ultimate goal of GWAS is to identify the underlying genes, which will in turn drive the introduction of new diversity into barley breeding programs via allele mining of ex situ primitive barley germplasm (Bockelman and Valkoun, 2010). Identification of major genes in barley by GWAS is now a reality (Cockram et al., 2010; Ramsay et al., 2011; Comadran et al., 2012) and we expect that this approach will succeed soon for QTL such as those described here. Increased links between genetic and genomic (sequence) resources (Schulte et al., 2009; Mayer et al., 2012) will increase the speed and efficiency of this approach.

Conclusions

Our study shows that many genomic regions of European two-row spring barley still show encouragingly high genetic diversity. Also, the GWAS resource described here has proven to be a promising tool, with broad application to the identification of site-specific genetic effects for agronomic traits (Tester and Langridge, 2010). These results hold great promise for future improvement of this gene pool via selective breeding and the markers described here will prove useful for such activities. However, three prominent regions of the spring two-row barley genome are approaching fixation in this germplasm. This regional extinction of diversity is disturbing, because it may correspond to major targets for modern barley breeding, raising the question whether “breeding the best with the best” will lead to a performance ceiling, as the best alleles available in an increasingly bottlenecked gene pool are combined together.

Supplemental Information Available

Supplemental material is available at <http://www.crops.org/publications/tpg>.

Supplemental Figure S1. Intrachromosomal linkage disequilibrium (LD) decay in the spring two-row barley cultivars.

Supplemental Figure S2. STRUCTURE (Pritchard et al., 2000, 2010) analysis of the spring two-row barley cultivars. Log probability data and ΔK plot.

Supplemental Figure S3. Estimation of multienvironmental variance components for the traits analyzed in this study. Trial data from 213 lines for a *Hordeum vulgare* spring two-row cultivar collection (HVCC₂₁₃) are included in this study.

Supplemental Table S1. List of accessions included in the *Hordeum vulgare* spring two-row cultivar collection (HVCC₂₁₆).

Supplemental Table S2. Trait variation for all the trait and environment combinations from the trials, plus optimal models used to correct for spatial trends in the data.

Acknowledgments

This work was supported by the ERA-PG-funded project Exbardiv (“Genomics-Assisted Analysis and Exploitation of Barley Diversity”; <http://www.erapg.org>) and BBSRC Project BB/I022899/1 “The diversity and evolution of the gene component of barley pericentromeric heterochromatin.” We thank the following persons for excellent technical assistance: Hazel Bull, Birgit Dubsy, Nadia Faccini, Heike Harms, Christiane Kehler, Richard Keith, Ute Krajewski, Leena Lohermaa, Marja-Leena Manninen, Marita Nix, Donata Pagani, Peter Schreiber, Kerstin Wolf, and the “Experimental Fields and Nurseries” and “Genome Diversity” group at IPK. We are also grateful to the following for support and discussions: Christine Hackett, Andreas Graner, Raj K. Pasam, Kerstin Neumann, and Michele Stanca.

References

- Barrett, J.C., B. Fry, J. Maller, and M.J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265. doi:10.1093/bioinformatics/bth457
- Bockelman, H.E., and J. Valkoun. 2010. Barley germplasm conservation and resources. In: S.E. Ullrich, editor, *Barley: Improvement, production, and uses*. Wiley-Blackwell, Oxford, UK. p. 144–159.
- Boer, M.P., D. Wright, L.Z. Feng, D.W. Podlich, L. Luo, M. Cooper, and F.A. van Eeuwijk. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. *Genetics* 177:1801–1813. doi:10.1534/genetics.107.071068
- Borovkova, I.G., B.J. Steffenson, Y. Jin, A. Kilian, A. Kleinhofs, and T.K. Blake. 1997. Identification and mapping of a leaf rust resistance gene in barley line Q21861. *Genome* 40:236–241. doi:10.1139/g97-033
- Botstein, D., R.L. White, M. Skolnick, and R.W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32:314–331.
- Burton, R.A., S.A. Jobling, A.J. Harvey, N.J. Shirley, D.E. Mather, A. Bacic, and G.B. Fincher. 2008. The genetics and transcriptional profiles of the cellulose synthase-like *HvCslF* gene family in barley. *Plant Physiol.* 146:1821–1833. doi:10.1104/pp.107.114694
- Christiansen, M.J., S.B. Andersen, and R. Ortiz. 2002. Diversity changes in an intensively bred wheat germplasm during the 20th century. *Mol. Breed.* 9:1–11. doi:10.1023/A:1019234323372
- Cleveland, D.A., and D. Soleri. 2002. Introduction. In: D.A. Cleveland and D. Soleri, editors, *Farmers, scientists and plant breeding: Integrating knowledge and practice*. CABI Publishing International, Wallingford, UK. p. 1–18.

- Close, T.J., P.R. Bhat, S. Lonardi, Y. Wu, N. Rostoks, L. Ramsay, et al. 2009. Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics* 10:582. doi:10.1186/1471-2164-10-582
- Cockram, J., J. White, D.L. Zuluaga, D. Smith, J. Comadran, M. Macaulay, et al. 2010. Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl. Acad. Sci. USA* 107:21611–21616. doi:10.1073/pnas.1010179107
- Comadran, J., B. Kilian, J. Russell, L. Ramsay, N. Stein, M. Ganal, et al. 2012. Natural variation in a homolog of *Antirrhinum* CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44:1388–1392. doi:10.1038/ng.2447
- Comadran, J., L. Ramsay, K. MacKenzie, P. Hayes, T.J. Close, G. Muehlbauer, et al. 2011. Patterns of polymorphism and linkage disequilibrium in cultivated barley. *Theor. Appl. Genet.* 122:523–531. doi:10.1007/s00122-010-1466-7
- Comadran, J., W.T.B. Thomas, F.A. van Eeuwijk, S. Ceccarelli, S. Grando, A.M. Stanca, et al. 2009. Patterns of genetic diversity and linkage disequilibrium in a highly structured *Hordeum vulgare* association-mapping population for the Mediterranean basin. *Theor. Appl. Genet.* 119:175–187. doi:10.1007/s00122-009-1027-0
- Davey, J.W., P.A. Hohenlohe, P.D. Etter, J.Q. Boone, J.M. Catchen, and M.L. Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12:499–510. doi:10.1038/nrg3012
- Donini, P., J.R. Law, R.M.D. Koebner, J.C. Reeves, and R.J. Cooke. 2000. Temporal trends in the diversity of UK wheat. *Theor. Appl. Genet.* 100:912–917. doi:10.1007/s001220051370
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14:2611–2620. doi:10.1111/j.1365-294X.2005.02553.x
- Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47–50.
- Fischbeck, G. 1992. Barley cultivar development in Europe – Success in the past and possible changes in the future. In: L. Munck, K. Kirkegaard, and B. Jensen, editors, *Barley genetics VI, Proceedings of the Sixth International Barley Genetics Symposium, Vol. II*, Helsingborg, Sweden. 22–27 July, 1991. Munksgaard International Publishers, Copenhagen, Denmark. p. 885–901.
- Gupta, P.K., S. Rustgi, and R.R. Mir. 2008. Array-based high-throughput DNA markers for crop improvement. *Heredity* 101:5–18. doi:10.1038/hdy.2008.35
- Haahr, V., and D. von Wettstein. 1976. Studies of an induced, high yielding dwarf-mutant of spring barley. In: H. Gaul, editor, *Barley genetics III, Proceedings of the Third International Barley Genetics Symposium, Garching, Germany. 7–12 July 1975*. Verlag Karl Thieme Publishers, Munich, Germany. p. 215–218.
- Hayes, P.M., B.H. Liu, S.J. Knapp, F. Chen, B. Jones, T. Blake, et al. 1993. Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. *Theor. Appl. Genet.* 87:392–401. doi:10.1007/BF01184929
- Inostroza, L., A. del Pozo, I. Matus, D. Castillo, P. Hayes, S. Machado, and A. Corey. 2009. Association mapping of plant height, yield, and yield stability in recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp. *vulgare* background. *Mol. Breed.* 23:365–376. doi:10.1007/s11032-008-9239-6
- Jing, R., A. Vershinin, J. Grzebyta, P. Shaw, P. Smykal, D. Marshall, M.J. Ambrose, et al. 2010. The genetic diversity and evolution of field pea (*Pisum*) studied by high throughput retrotransposon based insertion polymorphism (RBIP) marker analysis. *BMC Evol. Biol.* 10:44. doi:10.1186/1471-2148-10-44
- Khlestkina, E.K., X.Q. Huang, F.J.-B. Quenum, S. Chebotar, M.S. Röder, and A. Börner. 2004. Genetic diversity in cultivated plants—loss or stability? *Theor. Appl. Genet.* 108:1466–1472. doi:10.1007/s00122-003-1572-x
- Kilian, B., and A. Graner. 2012. NGS technologies for analyzing germplasm diversity in genebanks. *Briefings Funct. Genomics* 11:38–50. doi:10.1093/bfgp/eln046
- Kilian, B., H. Ozkan, J. Kohl, A. von Haeseler, F. Barale, O. Deuschet, et al. 2006. Haplotype structure at seven barley genes: Relevance to gene pool bottlenecks, phylogeny of ear type and site of barley domestication. *Mol. Genet. Genomics* 276:230–241. doi:10.1007/s00438-006-0136-6
- Kilian, B., H. Ozkan, A. Walther, J. Kohl, T. Dagan, F. Salamini, and W. Martin. 2007. Molecular diversity at 18 loci in 321 wild and 92 domesticated lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (Einkorn) domestication: Implications for the origin of agriculture. *Mol. Biol. Evol.* 24:2657–2668. doi:10.1093/molbev/msm192
- Koebner, R.M.D., P. Donini, J.C. Reeves, R.J. Cooke, and J.R. Law. 2003. Temporal flux in the morphological and molecular diversity of UK barley. *Theor. Appl. Genet.* 106:550–558.
- Kraakman, A.T.W., R.E. Niks, P. Van den Berg, P. Stam, and F.A. Van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168:435–446. doi:10.1534/genetics.104.026831
- Laidò, G., D. Barabaschi, A. Tondelli, A. Gianinetti, A.M. Stanca, O. Li Destri Nicosia, et al. 2009. QTL alleles from a winter feed type can improve malting quality in barley. *Plant Breed.* 128:598–605. doi:10.1111/j.1439-0523.2009.01636.x
- Lundqvist, U., J. Franckowiak, and T. Konishi. 1996. New and revised descriptions of barley genes. *Barley Genet. Newsl.* 26:34–43.
- Maccaferri, M., M.C. Sanguineti, P. Donini, and R. Tuberosa. 2003. Microsatellite analysis reveals a progressive widening of the genetic basis in the elite durum wheat germplasm. *Theor. Appl. Genet.* 107:783–797. doi:10.1007/s00122-003-1319-8
- Malosetti, M., F.A. van Eeuwijk, M.P. Boer, A.M. Casas, M. Elía, M. Moralejo, et al. 2011. Gene and QTL detection in a three-way barley cross under selection by a mixed model with kinship information using SNPs. *Theor. Appl. Genet.* 122:1605–1616. doi:10.1007/s00122-011-1558-z
- Malysheva-Otto, L., M.W. Ganal, J.R. Law, J.C. Reeves, and M.S. Röder. 2007. Temporal trends of genetic diversity in European barley cultivars (*Hordeum vulgare* L.). *Mol. Breed.* 20:309–322. doi:10.1007/s11032-007-9093-y
- Marquez-Cedillo, L.A., P.M. Hayes, A. Kleinhofs, W.G. Legge, B.G. Rossnagel, K. Sato, et al. 2001. QTL analysis of agronomic traits in barley based on the doubled haploid progeny of two elite North American varieties representing different germplasm groups. *Theor. Appl. Genet.* 103:625–637. doi:10.1007/PL00002919
- Mayer, K.F.X., R. Waugh, J.W.S. Brown, A. Schulman, P. Langridge, M. Platzer, et al. 2012. A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491:711–716.
- Moragues, M., J. Comadran, R. Waugh, I. Milne, A.J. Flavell, and J.R. Russell. 2010. Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. *Theor. Appl. Genet.* 120:1525–1534. doi:10.1007/s00122-010-1273-1
- Ordon, F., J. Ahlemeyer, K. Werner, W. Köhler, and W. Friedt. 2005. Molecular assessment of genetic diversity in winter barley and its use in breeding. *Euphytica* 146:21–28. doi:10.1007/s10681-005-5192-1
- Ovesná, J., L. Kučera, K. Vaculová, J. Milotová, J. Snape, P. Wenzl, et al. 2013. Analysis of the genetic structure of a barley collection using DNA Diversity Array Technology (DArT). *Plant Mol. Biol. Rep.* 31:280–288. doi:10.1007/s11105-012-0491-x
- Pasam, R.K., R. Sharma, M. Malosetti, F.A. van Eeuwijk, G. Haseneyer, B. Kilian, and A. Graner. 2012. Genome-wide association studies for agronomical traits in a worldwide spring barley collection. *BMC Plant Biol.* 12:16. doi:10.1186/1471-2229-12-16
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Pritchard, J.K., X. Wen, and D. Falush. 2010. Documentation for STRUCTURE software, version 2.3. University of Chicago, Chicago, IL. http://pritch.bsd.uchicago.edu/structure_software/release_versions/v2.3.4/structure_doc.pdf (accessed 2 Feb. 2010).

- Ramsay, L., J. Comadran, A. Druka, D.F. Marshall, W.T.B. Thomas, M. Macaulay, et al. 2011. INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. *Nat. Genet.* 43:169–172. doi:10.1038/ng.745
- Rosenberg, N.A., L. Huang, E.M. Jewett, Z.A. Szpiech, I. Jankovic, and M. Boehnke. 2010. Genome-wide association studies in diverse populations. *Nat. Rev. Genet.* 11:356–366. doi:10.1038/nrg2760
- Rostoks, N., L. Ramsay, K. MacKenzie, L. Cardle, P.R. Bhat, M.L. Roose, et al. 2006. Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proc. Natl. Acad. Sci. USA* 103:18656–18661. doi:10.1073/pnas.0606133103
- Russell, J., A. Booth, J. Fuller, B. Harrower, P. Hedley, G. Machray, and W. Powell. 2004. A comparison of sequence-based polymorphism and haplotype content in transcribed and anonymous regions of the barley genome. *Genome* 47:389–398. doi:10.1139/g03-125
- Russell, J., I.K. Dawson, A.J. Flavell, B. Steffenson, E. Weltzien, A. Booth, et al. 2011. Analysis of > 1000 single nucleotide polymorphisms in geographically matched samples of landrace and wild barley indicates secondary contact and chromosome-level differences in diversity around domestication genes. *New Phytol.* 191:564–578. doi:10.1111/j.1469-8137.2011.03704.x
- Russell, J.R., R.P. Ellis, W.T.B. Thomas, R. Waugh, J. Provan, A. Booth, et al. 2000. A retrospective analysis of spring barley germplasm development from ‘foundation genotypes’ to currently successful cultivars. *Mol. Breed.* 6:553–568. doi:10.1023/A:1011372312962
- Schulte, D., T.J. Close, A. Graner, P. Langridge, T. Matsumoto, G. Muehlbauer, et al. 2009. The international barley sequencing consortium – At the threshold of efficient access to the barley genome. *Plant Physiol.* 149:142–147. doi:10.1104/pp.108.128967
- Tester, M., and P. Langridge. 2010. Breeding technologies to increase crop production in a changing world. *Science* 327:818–822. doi:10.1126/science.1183700
- Thomas, W.T.B., W. Powell, R. Waugh, K.J. Chalmers, U.M. Barua, P. Jack, et al. 1995. Detection of quantitative trait loci for agronomic, yield, grain and disease characters in spring barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 91:1037–1047.
- Varshney, R.K., M.J. Paulob, S. Grando, F.A. van Eeuwijk, L.C.P. Keizer, P. Guo, et al. 2012. Genome wide association analyses for drought tolerance related traits in barley (*Hordeum vulgare* L.). *Field Crops Res.* 126:171–180. doi:10.1016/j.fcr.2011.10.008
- von Korff, M., H. Wang, J. Léon, and K. Pillen. 2006. AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (*H. vulgare* ssp. *spontaneum*). *Theor. Appl. Genet.* 112:1221–1231. doi:10.1007/s00122-006-0223-4
- VSN International. 2011. GenStat for Windows 14th edition. VSN International, Hemel Hempstead, UK.
- Wang, M., N. Jiang, T. Jia, L. Leach, J. Cockram, J. Comadran, et al. 2012. Genome-wide association mapping of agronomic and morphologic traits in highly structured populations of barley cultivars. *Theor. Appl. Genet.* 124:233–246. doi:10.1007/s00122-011-1697-2
- Yu, G.T., R.D. Horsley, B. Zhang, and J.D. Franckowiak. 2010. A new semi-dwarfing gene identified by molecular mapping of quantitative trait loci in barley. *Theor. Appl. Genet.* 120:853–861. doi:10.1007/s00122-009-1216-x
- Zhu, H., L. Gilchrist, P. Hayes, A. Kleinhofs, D. Kudrna, Z. Liu, et al. 1999. Does function follow form? Principal QTLs for Fusarium head blight (FHB) resistance are coincident with QTLs for inflorescence traits and plant height in a doubled-haploid population of barley. *Theor. Appl. Genet.* 99:1221–1232. doi:10.1007/s001220051328